An exploratory factor analysis of a survey intended to measure undergraduate student attitudes towards computational methods in physics

Andrew Gavrin[1], Greg Gallagher[1], Danka Maric[2], Gautam Vemuri[1]
[1]*Department of Physics, Indiana University Indianapolis,*
402 N Blackford St, Indianapolis, Indiana, 46202 and
[2]*STEM Education Innovation & Research Institute, Indiana University Indianapolis,*
755 W. Michigan Street, Indianapolis, Indiana, 46202

Computational methods of problem solving are increasingly emphasized by physics programs across the country, and some have adopted efforts to incorporate computational methods across the curriculum. However, there are no robust tools currently available that were designed to evaluate the effectiveness of these initiatives. This paper presents an exploratory factor analysis of data acquired using a survey developed to evaluate a department-wide computational initiative. The factor analysis supports previously published work establishing the validity and reliability of the survey, particularly a survey section which aims to gauge students' general attitudes toward using computational methods. The exploratory factor analysis conducted suggests a two-factor model, but with significant cross-loading between these factors for three of the survey items. This paper reports and interprets the results of the factor analysis and suggests improvements to the survey based on rephrasing the three cross-loaded items.

2024 PERC Proceedings edited by Ryan, Pawl, and Zwolak; Peer-reviewed, doi.org/10.1119/perc.2024.pr.Gavrin

I. INTRODUCTION

A. Integration of Computational Methods in Physics

The physics community has considered myriad ways to incorporate computers in undergraduate education over many decades [1], [2], [3], [4]. In recent years, the discussion has focused on teaching students to use computers as physicists do: for modeling complex problems, analyzing large data sets, producing static and animated visualizations, etc. [5–8]. Inclusion of computational thinking and methods across the curriculum has been promoted by the American Association of Physics Teachers (AAPT) [9], and by an APS-AAPT joint task force [10]. Most efforts to incorporate computation across the curriculum are in their early stages, with a few exceptions, e.g., Oregon State University [11] and Lawrence University [12].

Unfortunately, efforts to develop tools to evaluate these curricular efforts have lagged behind the efforts to teach computation. As of this writing, neither the PICUP project site [13] nor PhysPort [14] have suitable evaluation instruments available, but some efforts in this area are beginning [15]. Qualitative research on computational physics has shown that some faculty favor attitudinal measures as one means of assessing efforts to incorporate computational physics [16]. We have developed an instrument that evaluates some aspects of such a program, and reported on efforts to establish its reliability and validity [17]. This paper extends that work. A second paper, describing changes in students' self-efficacy with respect to computational methods, has also been submitted to PERC 2024. [18].

Section II of this paper will provide additional contextual information about our efforts and educational setting. Section III will outline our methodology for conducting the survey and conducting the exploratory factor analysis. Section IV will provide the results of these analyses, and Sections V and VI will present discussion and conclusions, including potential efforts to improve the survey based on the results obtained thus far.

II. CONTEXT

A. Institutional context

IU Indianapolis (formerly IUPUI) is an urban, public university, its Carnegie Classification is currently "R2: High Research Activity." [19]. Approximately 25,000 students are enrolled, of whom roughly 30% are people of color. The undergraduate population is approximately 17,000, and about 30% of that group are first generation college students. The undergraduate population is heavily weighted ($> 80\%$) towards in-state students, with a majority coming from within a 50 mile radius. The Physics department currently has 13 full-time faculty and offers B.S., M.S., and Ph.D. degrees. The undergraduate curriculum follows a traditional model, including a two-course introductory sequence, two upper-level labs, six required advanced courses, and a research capstone.

B. Prior work

The physics department at Indiana University Indianapolis began working towards a department-wide adoption of computational methods in 2018. Our overall goal is to "make computation normal," that is, students who complete the program should recognize computational approaches as a "normal" way to solve problems, and feel well prepared to do so. To this end, we have set a target that 25% of all work in the undergraduate physics curriculum be computational in nature. With no well-established tool available to evaluate this project, we developed the survey instrument discussed here in-house. We have previously reported details of the development process, and of efforts to establish the validity and reliability of the instrument [17]. For the purposes of this paper, it is necessary to note that the survey items were developed in a three step process:

1. Initial items were developed by the project team.
2. The items were subsequently refined in department meetings, with a goal of attaining consensus support from the faculty.
3. Some items were adjusted, and several more added, after consultation with outside experts from the PICUP collaboration.

This survey was first implemented at the conclusion of the fall semester of 2018, at the conclusion of step 2. It was given at the end of each semester for four semesters before step 3. The updated survey was then given at the end of each semester beginning in the fall of 2020 and continuing through the present. Data from the spring semester of 2024 are not included in the analysis presented here due to time constraints.

The survey has two parts, both focused on the affective domain. Part 1 is intended to determine the degree to which students' general attitudes about computational methods are similar to those of experts. It is composed of 9 Likert-scale items such as "Computational, experimental, and analytical ("pencil and paper math") methods are all necessary in the field of physics." Part 2 is intended to measure students' self-efficacy with respect to specific computational tools and methods, such as matrix operations and the use of LaTeX in scientific communication. This paper focuses on the first part of the survey - measuring student attitudes. In particular, we wish to focus on the project goal that our students graduate with "expert-like" attitudes regarding the use of computational methods and a degree of confidence with doing so. To this end, we structure our effort around two research questions.

C. Present goals

This paper seeks to expand on our prior work. In addition to including a more complete data set, we focus on answering the following research questions:

RQ1 How many reliable and interpretable factors are present in the survey?

RQ2 Does the interpretation of each factor present match the initial goals in the development of the survey instrument?

It is important to note that the factor analysis conducted in this paper focuses on just Part 1 of the survey (e.g., the Likert-scale attitude items), with the initial goals of Part 1 being to (a) measure student attitudes regarding computation in physics in general, and (b) measure students' overall self-efficacy with respect to using computational methods. While some of Part 1 is intended to measure student self-efficacy, it is important to note that it is distinguished from Part 2 of the survey (also measures self-efficacy) by the broader scope of the survey items. Part 1 is intended to measure student self-efficacy with regards to computational methods as a whole, while Part 2 measures self-efficacy for a narrower set of computational skills/methods.

III. METHODS

A. Survey method and participants

At the conclusion of each fall and spring semester, a link to the survey is sent to all students completing physics majors' courses by one of the authors (MD) who is not a member of the physics department. Students are informed that the survey is designed to be completed each semester, so they should do so even if they have completed it previously. We intentionally structured the survey this way in order to provide information about students' development over time. Students are not offered any incentives, either monetary or academic to complete the survey. Students are informed that their participation is optional and can be withdrawn at any time. Students are also informed that neither the fact of their participation nor specifics of their responses will be available to their professors at any time. We acknowledge that this mechanism can produce self-selection bias, and consider this in drawing conclusions from our results.

The introductory portion of the survey requests students' names and student identification numbers, and provides check boxes for students to indicate which physics courses they completed that semester. MD removes all identifying characteristics and assigns a research ID number that cannot be connected to individuals by any member of the physics department.

The next 9 items ask students the Likert-scale "attitudinal" prompts described above. The latter 22 items are intended to measure students self-efficacy with respect to a variety of computational skills and platforms. Results from those have been submitted in a separate manuscript [18]. For brevity in later discussions, Table I provides brief identifiers for each of the survey prompts analyzed in the factor analysis.

Only surveys completed from the fall semester of 2020 through the fall semester of 2023 are used here. Data from the first 4 semesters of the survey are excluded since several of the questions had changed slightly. We also exclude surveys that are only partially complete. Occasionally, a student will complete (or partially complete) the survey more than once in a single semester. In such cases, we retain only the last complete survey from that student in that semester. Finally, a Mahalanobis distance test was completed to check for outliers using recommended guidelines of significance ($p \leq 0.001$), and these outliers were discarded as well [20]. Overall, these requirements lead to dropping 24 surveys, including seven for incomplete responses, four for repeat surveys, and 13 outliers. The final data set analyzed for this work included $N = 356$ surveys, with $N_{100} = 180, N_{200} = 139, N_{300} = 23$, and $N_{400} = 14$. These levels roughly correspond to students in their first year (100-level) second year (200-level), etc. Students do occasionally take courses out of order, which we note is a source of noise in our data.

TABLE I. The survey items analyzed with convenient identifiers. Students are asked to provide their agreement with each item on a Likert-scale which is then converted to a numeric 1-5 scale for analysis.

Identifier	Survey Item
L1	Computational, experimental, and analytical ("pencil and paper math") methods are all necessary in the field of physics.
L2	Using computational methods helps me understand physics topics.
L3	I can sometimes use computational methods to understand problems that I am unable to do analytically.
L4	I can judge whether a given problem is most easily/better solved by computational vs. analytical means.
L5	Computational methods can be useful in conjunction with analytical methods for understanding physical phenomena.
L6	Using analytical and experimental methods helps me understand physics topics.
L7	I'm confident in the results of codes I develop.
L8	I feel well-prepared for using computational methods in graduate school or a future job.
L9	I have used computational methods outside my classes, such as in a research project, internship, or job.

B. Data analysis plan

To examine the psychometric properties of the nine survey items, we focused on assessing two sources: internal structure validity and internal consistency which we operationalize using the *Standards* set by a joint committee of education and psychology research organizations [21]. Internal structure validity, or the degree to which the relationship between items and components align with the latent construct that informs survey interpretation, was assessed using an exploratory factor analysis (EFA). Internal consistency, a form of reliability evidence that demonstrates the extent of agreement between items, was assessed using Cronbach's alpha coefficient.

IV. RESULTS

A. Preliminary analyses

As a preliminary assessment, we evaluated whether the nine items were appropriate for factor analysis based on the following criterion: (1) linearity between items, (2) factorability and multicollinearity, (3) normality, and (4) the absence of outliers.

Linearity was determined by examining scatterplots for all possible relationships between survey items. Upon examination, all item relationships appear to exhibit linear relationships. It is important to note that the size of the linear relationship is also important, as the relationships must be large enough to indicate factorability, but not so sizable as to be redundant. Factorability was determined by computing inter-item correlations, with results ranging from $|0.16|$ to $|0.63|$, with a majority (69%) showing at least a moderate correlation ($\geq |0.32|$) based on recommended ranges [20]. Further, none of the correlations fell above the recommended upper limit of $|0.70|$, indicating a lack of multicollinearity [22].

Univariate normality of the data was screened by examining the skewness and kurtosis values of each item. The skewness for all items fell below the recommended value for normality of $|2.0|$ [23]. The kurtosis values for most items fell below this threshold as well, except for one item which had a kurtosis of 2.26, though this value still falls well below the liberal kurtosis guideline of $|7.0|$, and thus was determined to not be of concern [22].

B. Psychometric Analyses

Before conducting the EFA, we used Horn's Parallel Analysis in STATA (version 17) to estimate an initial number of factors because mechanical rules of thumb (i.e., examining the scree plot or eigenvalues) alone do not always produce reliable factor structures [24]. The results of the parallel analysis suggest that two factors are present. The EFA was then conducted using a Principle Component Analysis (PCA), where a direct oblimin rotation was set to zero, the number of factors constrained was set to two, and all coefficients less than the recommended 0.32 moderate correlation were suppressed.

TABLE II. Results of the principle component analysis with factor loading and communalities (h^2) for survey items. Bold text indicates the sole or larger loading. Factor loadings displayed are based on the rotated structure matrix. The percent variance explained per factor (F_1 and F_2) is also included.

Survey Item	F_1	F_2	h^2
L1		**0.83**	0.71
L2	0.54	**0.65**	0.51
L3	0.55	**0.60**	0.47
L4	**0.69**	0.40	0.49
L5		**0.84**	0.71
L6	0.53	**0.60**	0.46
L7	**0.84**	0.42	0.71
L8	**0.82**		0.67
L9	**0.71**		0.52
% of variance explained	44.47	13.82	

Additionally, a Bartlett's test of sphericity and Kaiser-Meyer-Olkin (KMO) measure were conducted. Both are statistical tests meant to justify an EFA by determining whether measured variables are sufficiently intercorrelated, as a statistically significant Bartlett's test combined with a KMO value above the recommended value of 0.7 indicate appropriate factorability [25]. The Bartlett's test results of $\chi^2(36) = 1054.29$, $p < 0.001$, and the KMO measure result of KMO $= 0.86$ indicate sufficient intercorrelations to justify proceeding with a factor analysis [25]. The component structure of the PCA explained 58.29% of the variance, with the first factor containing four items, and the second containing five. The solution with all item communalities, loadings, and variances explained by components can be seen in Table II.

The internal consistency analysis results in a value of Cronbach's $\alpha = 0.77$ for both factors one and two. These values are generally taken to indicate a relatively high degree of consistency among items, especially when triangulated with the EFA results and conceptual knowledge. [26].

V. DISCUSSION

The observed factor loadings in Table II indicate all nine items are strong enough to be useful. A commonly used categorization is that items above 0.55 be considered "good,", above 0.63 is "very good," and above 0.71 "excellent" [20].

While a two-factor structure was suggested and retained, there is cross-loading between factors for five of the nine items. This indicates that those questions do not clearly fall into a single category. In two cases, (L4, and L7) one of the loadings is substantially larger than the other. In three (L2, L3, and L6) the loadings are nearly equal. Rather than discarding these items, as is sometimes done, it is our plan to refine the survey in an attempt to reduce cross-loading for a more cohesive survey instrument.

To refine the survey, it is first necessary to understand the two-factor structure and the source of the cross-loading. What follows is a hypothesis that can be tested by further psychometric development and analysis.

We begin by categorizing the items according to the factor loading we observe, then tentatively identify the structures that underlie the groups. Based on this, we can make changes to refine the structure and test those changes in subsequent runs of the survey. Our initial categorization is as follows:

- **F1** L8 and L9 are exclusively associated with factor 1, and L4 and L7, despite some cross-loading, can be clearly identified with this factor.
- **F2** L1 and L5 are exclusively associated with factor 2.
- **M** (Mixed) L2, L3, and L6 have substantial cross-loading. The EFA places them more strongly with factor 2, but clearly improvement is needed.

We observe that all four items in **F1** begin with the word "I" and indicate significant confidence in a specific ability, e.g., L4, which is "I can judge whether a problem is most easily/better solved by computational vs. analytical means." As such, this factor appears to serve our goal of measuring students' overall comfort with computational methods.

We also note that items L1 and L5 make no reference to students' judgements of their own abilities. Rather, they are strictly focused on students' sense of what is generally true in physics as a discipline. Item L1 is a clear example: "Computational, experimental, and analytical ("pencil and paper math") methods are all necessary in the field of physics." This identification of **F2** corresponds to our goal of determining whether students have adopted "expert-like" opinions regarding the value of computational methods.

For the present two-factor structure, we tentatively have named the first factor, **F1**, "Self-Efficacy in Computational Physics." Similarly, we consider the second factor, **F2**, as "Attitudes towards Computational Physics."

We have labeled Items L2, L3, and L6 as **M** (mixed) due to the substantial cross-loading observed. At first glance, these items appear to be a better fit to **F1**, as they refer to the students' ability. However, we note that these items are softer in tone. They do not begin with "I", nor do they express complete confidence. Instead, they contain qualifiers, such as "sometimes" as in L3, which reads "I can sometimes use computational methods to understand problems that I am unable to do analytically." Similarly, items L2 and L6 both use the term "helps" as in "Using computational methods helps me understand physics topics." This is clearly a weaker statement than saying, e.g., "I can use computational methods to understand physics topics."

Based on this understanding of the factor analysis, we plan to rephrase the three mixed items. Our goal is a robust instrument with two clear factors; **F1** "self-efficacy" and **F2** "expert attitudes," as described above. At present, **F2** has only two strong items, so we will revise currently cross-loaded questions as below, with the intent to shift them clearly into **F2** by removing references to the students' own abilities.

- L2 (revised): Physicists can use computational methods to better understand physics topics.
- L3 (revised): Computational methods can be used to understand problems that cannot be solved analytically.
- L6 (revised): Physicists can use experimental and computational methods in a complementary way to understand complex problems.

VI. CONCLUSIONS

This paper sought to address the following two research questions:

- **RQ1** How many reliable and interpretable factors are present in the survey?
- **RQ2** Does the interpretation of each factor present match the initial goals in the development of the survey instrument?

Preliminary tests showed linear relationships among the nine survey items analyzed in this paper. An estimate of the number of factors in the survey using Horn's Parallel Analysis suggested two factors. An exploratory factor analysis was conducted using a Principle Component Analysis, and the results allow us to answer **RQ1** by identifying two reliable and interpretable factors present in the survey instrument.

The two factors, **F1** and **F2**, present were tentatively identified as "Self-Efficacy in Computational Physics" and "Attitudes towards Computational Physics," respectively. This identification of the factors provides a positive answer to **RQ2**. The factors are well aligned with the goals of the instrument as an evaluation tool for our departmental initiative.

However, there is significant cross-loading between the factors for three survey items, weakening the identification of the two factors. Based on this, we plan to re-structure the survey by refining three of the cross-loaded items so they fit better with **F2**. Subsequent runs of the survey will determine whether this effort is successful.

We recognize that our instrument needs refining, but believe this can be accomplished easily. As such, we believe it can provide a useful step towards creating a robust set of computational physics assessments that can be used by the community.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the STEM Education and Innovation Research Institute (SEIRI) at IU Indianapolis, and from National Science Foundation grant DUE-2021209. We also wish to thank the Partnership for Integration of Computation into Undergraduate Physics (PICUP).

[1] J. M. Wilson and E. F. Redish, Using Computers in Teaching Physics, Physics Today **42**, 34 (1989).
[2] E. F. Redish, Is the Computer Appropriate for Teaching Physics?, Computers in Physics **7**, 613 (1993).
[3] R. H. Landau, Resource Letter CP-2: Computational Physics, American Journal of Physics **76**, 296 (2008).
[4] T. J. Atherton, Resource Letter CP-3: Computational physics, American Journal of Physics **91**, 7 (2023).
[5] J. R. Taylor and B. A. King, Using computational methods to reinvigorate an undergraduate physics curriculum, Comput. in Sci. and Eng. **8**, 38 (2006).
[6] R. S. Chabay and B. A. Sherwood, Computational physics in the introductory calculus-based course, Am. J. Phys. **76**, 307 (2008).
[7] R. M. Serbanescu, P. J. Kushner, and S. Stanley, Putting computation on a par with experiments and theory in the undergraduate physics curriculum, Am. J. Phys. **79**, 919 (2011).
[8] C. M. Orban and R. M. Teeling-Smith, Computational Thinking in Introductory Physics, The Physics Teacher **58**, 247 (2020).
[9] AAPT recommendations for computational physics in the undergraduate physics curriculum, https://www.aapt.org/Resources/upload/AAPT_UCTF_CompPhysReport_final_B.pdf (2016), retrieved 5/16/21.
[10] P. Heron and L. McNeil, *PHYS21: Preparing Physics Students for 21st Century Careers* (American Physical Society, College Park, MD, 2016).
[11] D. H. McIntyre, J. Tate, and C. A. Manogue, Integrating computational activities into the upper-level Paradigms in Physics curriculum at Oregon State University, Am. J. Phys. **76**, 340 (2008).
[12] D. M. Cook, Computation in the Lawrence physics curriculum, https://www2.lawrence.edu/dept/physics/ccli/compatlu.pdf (2006), retrieved 5/16/21.
[13] Partnership for the Integration of Computation into Undergraduate Physics, https://www.compadre.org/picup/, retrieved 5/25/21.
[14] PhysPort, https://www.physport.org/, retrieved 6/1/21.
[15] M. D. Caballero, M. A. Kohlmyer, and M. F. Schatz, Implementing and assessing computational modeling in introductory mechanics, Phys. Rev. ST Phys. Educ. Res. **8**, 020106 (2012).
[16] J. Gambrell and E. Brewe, Towards a generalized assessment of computational thinking for introductory physics students (2023), arXiv:2308.03593 [physics.ed-ph].
[17] A. Gavrin, G. Vemuri, and D. Maric, Preliminary efforts to evaluate an initiative introducing computation across the undergraduate physics curriculum, 2021 PERC Proceedings (2021).
[18] G. Gallagher, G. Vemuri, A. Gavrin, and D. Maric, Evaluating students' self-efficacy with respect to computational methods in the undergraduate physics curriculum, 2024 PERC Proceedings (submitted) (2024).
[19] Carnegie Classification of Institutions of Higher Education, https://carnegieclassifications.acenet.edu/institution/indiana-university-purdue-university-indianapolis//, retrieved 4/29/24.
[20] L. Tabachnik, B.G. & Fidell, *Using Multivariate Statistics* (Pearson Education Limited, 2012).
[21] *Standards for Educational and Psychological Testing* (American Educational Research Association, 2014).
[22] E. Kneckta, C. Runyon, and S. Eddy, One size doesn't fit all: Using factor analysis to gather validity evidence when using surveys in your research, CBE Life Sciences Education (2019).
[23] D. L. Bandalos and S. J. Finney, Factor analysis exploratory and confirmatory, in *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (Routledge, 2018).
[24] B. P. O'Connor, Spss and sas programs for determining the number of components using parallel analysis and velicer's map test, Behavior Research Methods, Instruments, & Computers (2000).
[25] M. W. Watkins, Eploratory factor analysis: A guide to best practice, Journal of Black Pscyhology (2018).
[26] K. S. Taber, The use of cronbach's alpha when developing and reporting research instruments in science education, Res Sci Educ (2018).